



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/11, 15/10, C07K 14/47, C12P 21/00, C12Q 1/68, C07K 16/18, G06F 17/30, 17/50	A2	(11) International Publication Number: WO 99/53051 (43) International Publication Date: 21 October 1999 (21.10.99)
--	----	--

(21) International Application Number: PCT/IB99/00712
(22) International Filing Date: 9 April 1999 (09.04.99)

(30) Priority Data:
09/057,719 9 April 1998 (09.04.98) US
09/069,047 28 April 1998 (28.04.98) US

(71) Applicant (for all designated States except US): GENSET
[FR/FR]; 24, rue Royale, F-75008 Paris (FR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): DUMAS MILNE ED-
WARDS, Jean-Baptiste [FR/FR]; 8, rue Grégoire-de-Tours,
F-75006 Paris (FR). DUCLERT, Aymeric [FR/FR]; 6 ter,
rue Victorine, F-94100 Saint-Maur (FR). GIORDANO,
Jean-Yves [FR/FR]; 12, rue Duhesme, F-75018 Paris (FR).

(74) Agents: MARTIN, Jean-Jacques et al.; Cabinet Regimbeau,
26, avenue Kléber, F-75116 Paris (FR).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE,
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,
NL, PT, SE).

Published

*Without international search report and to be republished
upon receipt of that report.*

(54) Title: 5' ESTS AND ENCODED HUMAN PROTEINS

(57) Abstract

The sequences of 5' ESTs derived from mRNAs encoding secreted proteins are disclosed. The 5' ESTs may be to obtain cDNAs and genomic DNAs corresponding to the 5' ESTs. The 5' ESTs may also be used in diagnostic, forensic, gene therapy, and chromosome mapping procedures. Upstream regulatory sequences may also be obtained using the 5' ESTs. The 5' ESTs may also be used to design expression vectors and secretion vectors.

A6 - 09/563,516

identify a sequence encoding a signal peptide in these nucleic acids. The locations of the complete ORFs are listed in the accompanying Sequence Listing.

SEQ ID NOs. 812-1516 are "incomplete polypeptide sequences" which include a signal peptide. "Incomplete polypeptide sequences" are polypeptide sequences encoded by nucleic acids in which a start
5 codon has been identified but no stop codon has been identified. These polypeptides are encoded by the nucleic acids of SEQ ID NOs. 24-728. The location of the signal peptide, the von Heijne score of the signal peptide, the sequence of the signal-peptide and the proteolytic cleavage site are indicated as described above.

SEQ ID NOs. 1517-1553 are incomplete polypeptide sequences in which no signal peptide has
10 been identified to date. However, it remains possible that subsequent analysis will identify a signal peptide in these polypeptides. These polypeptides are encoded by the nucleic acids of SEQ ID NOs. 729-765.

SEQ ID NOs. 1554-1580 are "complete polypeptide sequences" which include a signal peptide. "Complete polypeptide sequences" are polypeptide sequences encoded by nucleic acids in which a start
15 codon and a stop codon have been identified. These polypeptides are encoded by the nucleic acids of SEQ ID NOs. 766-792. The location of the signal peptide, the von Heijne score of the signal peptide, the sequence of the signal-peptide and the proteolytic cleavage site are indicated as described above..

SEQ ID NOs. 1581-1599 are complete polypeptide sequences in which no signal peptide has been identified to date. However, it remains possible that subsequent analysis will identify a signal
20 peptide in these polypeptides. These polypeptides are encoded by the nucleic acids of SEQ ID NOs. 793-811.

SEQ ID NOs. 1600-1622 are nucleic acid sequences in which no open reading frame has been conclusively identified to date. However, it remains possible subsequent analysis will identify an open reading frame in these nucleic acids.

25 In the accompanying Sequence Listing, all instances of the symbol "n" in the nucleic acid sequences mean that the nucleotide can be adenine, guanine, cytosine or thymine. In some instances the polypeptide sequences in the Sequence Listing contain the symbol "Xaa." These "Xaa" symbols indicate either (1) a residue which cannot be identified because of nucleotide sequence ambiguity or (2) a stop codon in the determined sequence where applicants believe one should not exist (if the sequence
30 were determined more accurately). In some instances, several possible identities of the unknown amino acids may be suggested by the genetic code.

In the case of secreted proteins, it should be noted that, in accordance with the regulations governing Sequence Listings, in the appended Sequence Listing, the full protein (*i.e.* the protein containing the signal peptide and the mature protein) extends from an amino acid residue having a
35 negative number through a positively numbered C-terminal amino acid residue. Thus, the first amino acid of the mature protein resulting from cleavage of the signal peptide is designated as amino acid

C) Deletions in the sequence of a consensus contigated 5'EST to derive a preferred nucleic acid fragment are denoted by an "D", followed by a number indicating the first nucleotide position in a specific SEQ ID to be deleted in a string of deleted nucleotides or the position of the deleted nucleotide in the case of a single deleted nucleotide. Then there is a comma followed by number indicating the number of nucleotide(s) deleted from the sequence provided in the sequence ID. For example, SEQ ID NO: 5398; Position of preferred fragments: 56-780; Variant nucleotides D114,5 would indicate that a preferred polynucleotide fragment had the sequence of positions 56 to 780 of SEQ ID NO. 5398, except that the nucleotides in positions 114 to 118 had been deleted in the preferred polynucleotide as compared with the sequence of SEQ ID No. 5398.

The present invention encompasses isolated, purified, or recombinant nucleic acids which consist of, consist essentially of, or comprise a contiguous span of at least 8, 10, 12, 15, 18, 20, 25, 35, 40, 50, 70, 80, 100, 250, or 500 nucleotides in length, to the extent that a contiguous span of these lengths is consistent with the lengths of the particular polynucleotide, of a polynucleotide described in Table II, or a sequence complementary thereto, wherein said polynucleotide described in Table II is selected individually or in any combination from the polynucleotides described in Table II. The present invention also encompasses isolated, purified, or recombinant nucleic acids which consist of or consist essentially of a polynucleotide described in Table II, or a sequence complementary thereto, wherein said polynucleotide is selected individually or in any combination from the polynucleotides described in Table II. The present invention further encompasses isolated or purified polypeptides which consist of, consist essentially of, or comprise a contiguous span of at least 8, 10, 12, 15, 18, 20, 25, 35, 40, 50, 70, 80, or 100 amino acids encoded by a polynucleotide described in Table II.

Table II

SEQ ID NO.	Positions of Preferred Fragments	Variant nucleotides
35	1-423	S124, s; I135, a; S293, w; I363, a; S377, r; D424, 15
41	1-427	I117, m; S120, r; S124, g; D373, l; S376, b; S378, b; I427, gggg; D428, 109
43	1-276	S114, m; S118, rg; S123, r; S139, nr; I142, t; D148, l; D152, l; I228, t; I276, gg; D277, 136
45	126-420	D1, 125; I420, ggg; D421, 100
46	1-255	S139, r; I145, r; S146, mm; S150, ar; S254, g; D256, 128
48	4-437	D1, 3; S49, a; S55, g; S79, a; S90, a; I437, tctctg
59	1-471	S26, a; S44, t; S48, t; S109, a; S191, t; S200, gc; S203, a; S210, g; S237, a; S240, g; S255, a; S272, a; S277, a; S279, a; S284, t; S297, g; S305, g; S316, a; I471, ggtca
66	1-428	I428, tactgggg

		S338, c; I341, cccccggg; D342, 218
805	2-409	D1, 1; S334, d; I409, aggg; D410, 161
806	5-384	D1, 4; I384, actaa
807	1-301	S113, a; S117, c; S123, t; D128, 1; D134, 1; S282, g; S284, a; I301, gacggagggg; D302, 70
808	2-314	D1, 1; S306, g; I314, ggg; D315, 121
809	1-394	S53, g; S228, n; S272, vk; I301, g; I358, m; S368, nb; S375, w; I383, mm; I388, yt; I394, nhaccggg
810	6-205	I0, a; D1, 5; I141, t; I205, ggg; D206, 630
811	6-270	D1, 5; I270, gggg; D271, 115
1600	1-247	S45, m; S114, k; I122, m; S123, yc; S158, rr; S221, k; I247, ccccaggg
1601	1-225	S109, bm; S195, m; I225, tgcacggg
1602	23-245	D1, 22; D138, 1; S139, s; S242, t; S244, g; I245, g; D246, 13
1603	1-303	S71, c; D277, 1; I303, ggagggg; D304, 38
1604	1-242	S47, w; S50, c; S81, h; S85, d; S91, k; S106, r; I242, tgtggg; D243, 50
1605	2-225	D1, 1; S20, k; S91, c; I225, ggg; D226, 132
1606	15-293	D1, 14; S156, g; S193, g; I200, t; I293, acaaaggg
1607	1-361	S323, c; I361, cccca
1608	1-151	I151, taagggg; D152, 154
1609	1-242	S55, s; I135, a; S152, h; I242, cagtaggg
1610	1-196	I151, w; S190, k; I196, cctgtgg
1611	1-228	S115, k; S174, rk; I228, cgtttggg
1612	1-221	S108, v; I221, tgatcggg
1613	1-281	I66, w; I137, a; D282, 79
1614	1-171	S53, k; S76, k; I80, k; S81, kw; S86, r; S92, k; S126, k; I171, gccgagg
1615	2-193	D1, 1; S67, c; I121, s; S122, mm; S126, g; S130, r; S146, r; S156, gm; I193, cctca
1616	1-349	S251, ww; S259, rs; S275, k; I279, w; S285, y; S292, y; I320, m; I331, m; I338, w; I341, s; I349, accccggg
1617	1-129	I118, t; D130, 26
1618	1-184	D9, 1; D185, 1
1619	1-169	I122, t; I169, gccaggg
1620	1-187	S106, k; S118, m; S122, cg; S132, k; D188, 59
1621	1-153	D125, 1; I131, ttt; S152, t; I153, gg; D154, 127
1622	1-400	S43, s; I126, g; I129, y; S353, d; I400, tatat

EXAMPLE 16

Categorization of 5' ESTs and Consensus Contigated 5'ESTs

The nucleic acid sequences of the present invention (SEQ ID NOs. 24-811 and 1600-1622) were
 5 grouped based on their homology to known sequences as follows. All sequences were compared to
 EMBL release 57 and daily releases available at the time of filing using BLASTN. All matches with a
 minimum of 25 nucleotides with 90% homology were retrieved and used to compute Tables IV and V.

In some embodiments, 5'ESTs or consensus contigated 5'ESTs nucleic acid sequence do not match any known vertebrate sequence nor any publicly available EST sequence, thus being completely new.

In other embodiments, 5'ESTs or consensus contigated 5'ESTs match a known sequence.

- 5 Tables III and IV gives for each sequence of the invention in this category referred to by its sequence identification number in the first column, the positions of their preferred fragments in the second column entitled "Positions of preferred fragments." As used herein the term "polynucleotide described in Table III" refers to the all of the preferred polynucleotide fragments defined in Table III in this manner, and the term "polynucleotide described in Table IV" refers to the all of the preferred polynucleotides fragments
- 10 defined in Table IV in this manner. The present invention encompasses isolated, purified, or recombinant nucleic acids which consist of, consist essentially of, or comprise a contiguous span of at least 8, 10, 12, 15, 18, 20, 25, 35, 40, 50, 70, 80, 100, 250, or 500 nucleotides in length, to the extent that a contiguous span of these lengths is consistent with the lengths of the particular polynucleotide, of a polynucleotide described in Table III or Table IV, or a sequence complementary thereto, wherein said
- 15 polynucleotide described in Table III or Table IV is selected individually or in any combination from the polynucleotides described in Table III or Table IV. The present invention also encompasses isolated, purified, or recombinant nucleic acids which consist of or consist essentially of a polynucleotide described in Table III or Table IV, or a sequence complementary thereto, wherein said polynucleotide is selected individually or in any combination from the polynucleotides described in Table III or Table IV.

20

Table III

SEQ ID NO	Positions of preferred fragments
24	1-251
25	1-83
28	227-276
29	1-27
30	130-242, 283-315, 365-461
32	314-399
33	89-321
34	1-38
35	1-52, 171-222
36	1-30, 408-441
37	1-138
39	115-140
40	1-97
41	1-112
42	1-177
46	1-38
48	376-400
51	400-466
54	1-259
55	189-320

742	217-280
743	10-275
747	1-179
749	2-31, 139-168
750	349-410
752	1-119
753	1-121
754	1-28
760	25-175
761	1-212
763	8-75
766	1-59, 102-248, 295-320
769	53-85
771	1-370
774	1-347
776	1-200
778	39-342
779	4-28
780	1-49, 407-472
781	116-426
782	1-59
783	1-53, 219-453
784	29-53, 219-263, 426-494
785	99-347, 386-461
786	2-28
788	1-279
789	1-58
790	226-268
792	129-218
794	265-431
796	5-86
797	1-34
799	1-344
802	46-477
806	64-384
807	135-301
808	2-314
810	6-39
1600	1-25
1601	1-225
1602	23-139
1603	1-294
1606	15-44
1607	1-361
1611	85-228
1612	1-221
1613	138-281
1614	65-171
1615	2-142
1616	1-46
1617	1-95
1620	1-187
1621	1-136

Table V

SEQ ID NO	Tissue Distribution
24	AA:1
25	S:1
26	P:1
27	W:1
28	P:1
29	S:1
30	P:1
31	P:1
32	P:1
33	P:1
34	AB:1
35	G:3; P:1; S:1; W:3; AA:4
36	P:1
37	S:1
38	Q:1
39	P:1
40	AB:1
41	B:1; C:3; F:1; G:1; H:4; S:2; T:8; W:1; Z:1; AA:3; AC:1; AD:3
42	A:1
43	N:2
44	P:1
45	C:2; K:1; O:1; S:5
46	K:1; S:2; AA:1
47	AA:1
48	C:1; O:1; P:8
49	P:1
50	P:1
51	P:1
52	S:1
53	AA:1
54	T:1
55	P:1
56	P:1
57	P:1
58	P:1
59	P:7; T:2; Z:1
60	R:1
61	C:1
62	P:1
63	F:1
64	AA:1
65	F:1

1615	AA:10
1616	B:4
1617	T:2
1618	K:2; S:8; AA:1
1619	B:2
1620	W:2
1621	H:1; AB:1
1622	H:2

Table VI

Tissue code	Tissue type
A	Bone Marrow
B	Brain
c	Cancerous prostate
D	Cerebellum
E	Colon
F	Dystrophic muscle
G	Fetal brain
H	Fetal kidney
I	Fetal liver
J	Heart
K	Hypertrophic prostate
L	Kidney
M	Large intestine
N	Liver
O	Lung
P	Lymph ganglia
Q	Lymphocytes
R	Muscle
S	Prostate
T	Ovary
U	Pancreas
V	Placenta
W	Spinal cord
X	Spleen
Y	Substantia nigra
Z	Surrenals
AA	Testis
AB	Thyroid
AC	Umbilical cord
AD	Uterus

- 5 In addition to categorizing the 5' ESTs and consensus contigated 5' ESTs with respect to their tissue of origin, the spatial and temporal expression patterns of the mRNAs corresponding to the 5' ESTs and consensus contigated 5' ESTs, as well as their expression levels, may be determined as described in Example 18 below.

atatcaagca aagaaaatgc caattctcag ccaaatgatg aagatgcctc ctctgatgcc 180
tactgctttg agctgctctc tatgggtttta gcaactgagt gctctaactg tggccggcaa 240
tatctggctc aacagctaac cctgcttcag gatctcttcc gctgcttcac acagcctctc 300
ctagagtcca gagacaggta cctcttttact agaagagttt gctgaagta 349

<210> 1617

<211> 155

<212> DNA

<213> Homo sapiens

<400> 1617

atacacatat ccatgggtttg tgagaggctc ctactaccc gtctgtctc agaatgtcag 60
aatgccctgt ttccttcctt tttgtggaca agtcaactct atacaatatt tgaagggatt 120
attctgaacc catctgaatg accaaggcct gaggc 155

<210> 1618

<211> 185

<212> DNA

<213> Homo sapiens

<400> 1618

cttgaatgg gctgagtcctc tcttgcctac ccttgacttg gaaaaaccag tttctctttt 60
attgtctgtt actaatctct attctaaaaa ttcagctcaa ttctcaacca tactccaaac 120
tctctctttt ccagctacct ttactcctc tccttcaatt ccactttcct ctgcttactt 180
ttttt 185

<210> 1619

<211> 169

<212> DNA

<213> Homo sapiens

<400> 1619

gggcgcaatg gcggatacgc tggagtcctc gctggaggac ccactgcgga gtttgtgcga 60
gttttggaga agcgggatgg tacagtgcct cgactacagc agtatagctc cggtggcgctg 120
ggtgcgttgt gtgggacgct gccattgtcc tttctaaata cctggaaac 169

<210> 1620

<211> 246

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> 122

<223> n=a, g, c or t

<400> 1620

cagaggtttt gttttcttca taatttttat cactatctga attattttga ttctttgttt 60
atttgtgcat ttcacgttgt ttcttatatt ccgttcaatg taagctctat gagaccaaga 120
anstgggcag ttttattcac cataagtatt ccaagcccta gtggttcctg gcacattttg 180
tattcacaat aaatatttgt taagtcaatg accagatgaa tggcttttaa actcaagata 240
gttttt 246

<210> 1621

<211> 280

<212> DNA

<213> Homo sapiens

<400> 1621

agtctaggga aagtcattca gtggatgtga tcttggtctc caggggacga tgtcaagctc 60
ttcttgctc cttctcagcc ttgttgctgt aactgctgct cagtccacca ttgaggaaca 120
ggccaaagac atttttgac aagtttaacc acgaagccga agactgttc tatcaaagtt 180
cacttgcttc ttggaattat aacaccaata ttactgaaga gaatgtccaa aacatgaata 240
atgctgggga caaatggtct gccttttttaa aggaacagtc 280

<210> 1622

INC305090

TCAAAGCATCAAAGTGAGGATAAGCCTAAAATCAGCTCTTGGAGATAAAGCA
TATGANTGGAACGACAATGAAATGTACCTGTTCCGATCATNTGTTGCATATGC
TATGAGGCAGTACTTTTTAAAAGTAAAAAATCAGATGATTCTTTTGGGGAG
GAGGATGTGCGAGTGGCTAATTTGAAACCAAGAATCTCCTTTAATTTCTTTGT
CACTGCACCTAAAAATGTGTCTGATATCATTCTAGAACTGAAGTTGAAAAG
GCCATCAGGATGTCCCGGAGCCGTATCAATGATGCTTTCCGTCTGAATGACA
ACAGCCTAGAGT

INC761848

GGAGTGATAGTGGTTGGCATTGTCATCCGATCTTCATGGGATCAGAGATCGG
AAGAAGAAAAATAAAGCAAGAAGTGGAGAAAATCCTATGCCTCCATCGTATT
AGCAAAGGAGAAATAATCCAGGTTCCAAAACACTGATGTGTTCAGACCTCCT
TTTAGAAAAATTATGTTTTTNCTTTGAGGTGTTTTGTTT

INC1426357

CTAAATGTAAATGTCTGTTGAATTTCTGAAGTTGAAAACAAGGATATATCATT
GGAGCAAGTGTTGGATCTTGTATGGAATATGGATGGATCACTTGTAAGGACA
GTGCCTGGGAACTGGTGTAGCTGCAAGGATTGAGAATGGCATGCATTAGCTC
ACTTTCATTTAATCCATTGTCAAGGATGACATGCTT

INC1501621

CTGCACCTAAAAATGTGTCTGATATCATTCCTAGAACTGAAGTTGAAAAGGC
CATCNGGATGTCCCGGAGCCGTATCAATGATGCTTTCCGTCTGAATGACAAC
AGCCTAGAGTTTCTGGGGATACAGCCAACACTTGGACCTCCTAACCAGCCCC
CTGTTTCCATATGGCTGATTGTTTTTGGAG

INC1582746

GTCTAAAGAGAGAAGAATCCAGGGAACAGGTAGAGGACATTGCTTTTTTCACT
TCCAAGGTGCTTGATCAACATCTCCCTGACAACACAAAACCTAGAGCCAGGGG
CCTCCGTGAACTCCCAGAGCATGCCTGATAGAACTCATTTCTACTGTTCTCT
AACTGTGGAGTGAATGGAAATTCCAACCTGTATGTTCACCCCTCTGAAGTGGGT
ACCCA

INC2328218

ANATGATAAAGATATCATTAAATGTCAAACTATGACTCTGTTTCAGAAAAAA
AATTGTCCAAAGACAACATGGCCAAGGAGAGAGCATCTTCATTGACATTGCT
TTCAGTATTTATTTCTGTCTCTGGATTGACTTCTGTTCTGTTTCTTAATAAGG
ATTTGTATTAGAGTATATTAGGGAAAGTGTGTATTTGGTCTCACAGGCTGTT
CAGGGATAATCTAAATGTAAATGTCTGTTGAATTCTGAAGTTGAAAACAA

INC2770719

CATGCTTTCTTCACAGTAACTCAGTTCAAGTACTATGGTGATTTGCCTACAGT
GATGTTTGGGAATCGATCATGCTTTCTTCAAGGTGACAGGTCTAAAGAGAGAA
GAATCCAGGGAACAGGTAGAGGACATTGCTTTTTCACTTCCAAGGTGCTTGA
TCAACATCTCCCTGACAACACAAAACACTAGAGCCAGGGGCCTCCGTGAACTCC
CAGAGCATGCCTGATAGAACTCATTCTACTGTTCTC

A13 - 09/563,516

INC2925464

TGATCTTGGCTCACAGGGGACGATGTCAAGCTCTTCCTGGCTCCTTCTCAGCC
TTGTTGNTGTAACTGCTGCTCAGTCCACCATTGAGGAACAGGCCAAGACATTT
TTGGACAAGTTTAACCACGAAGCCGAAGACCTGTTCTATCAAAGTTCACTTGC
TTCTTGGAATTATAACACCANTATTACTGAAGAGAATGTCCAAAACATGAAT
AATGCTGGGGACAAATGGTCTGCCTTTTTAAAGG

INC3279857

CATGAAGGCCCTCTGCACAAATGTGACATCTCAAACCTCTACAGAAGCTGGAC
AGAAACTGTTCAATATGCTGAGGCTTGGAAAATCAGAACCCTGGACCCTAGC
ATTGGAAAATGTTGTAGGAGCAAAGAACATGAATGTAAGGCCACTGCTCAAC
TACTTTGAGCCCTTATTTACCTGGCTGAAAGACCAGAACAAGAATTCTTTTGT
GGGATGGAGTACCGACTGGAGTCCATATGCAGACCAAAGCATCAAAGTGAG
GATAA

INC3222459

TGTCATCCTGATCTTCACTGGGATCAGAGATCGGAAGAAGAAAAATAAAGCA
AGAAGTGGAGAAAATCCTTATGCCTCCATCGATATTAGCAAAGGAGAAAATA
ATCCAGGATTCCAAAACACTGATGATGTTTCAGACCTCCTTTTAGAAAAATCTA
TGTTTTTCCTCTTGAGGTGATTTTGTTGTATGTAAATGTTAATTTTCATGGTATA
GAAATATAAGATGATAAAGATATCATTAAATGTCAAACTATGACTCTGTT
CAGAAAAAAAATTGTCCAAAGACAACATGGCCAAGGAGAGAGCATCTTCATT
GACATTGCTTTCAGTATTTATTTCTGTCTCTGGATT

INC4172634

AGAAAAATAAAGCAAGAAGTGGAGAAAATCCTTATGCCTCCATCGATATTAG
CAAAGGAGAAAATAATCCAGGATTCCAAAACACTGATGATGTTTCAGACCTCC
TTTAGAAAAATCTATGTTTTTCCTCTTGAGGTGATTTTGTTGTATGTAAATGT
TAATTTTCATGGTATAGAAAATATAAGATGATAAAGATATCATTAAATGTCAA
AACTATGACTCTGTTTCAGAAAAAAATTGTCCAAAGACAACATGGCCAAGGAG
AGAGCATC

INC4179741

ATTCAGTGGATGTGATCTTGGCTCACAGGGGACGATGTCAAGCTCTTCCTGGC
TCCTTCTCAGCCTTGTTGCTGTAAGTCTGCTCAGTCCACCATTGAGGAACAG
GCCAAGACATTTTTGGACAAGTTTAACCACGAAGCCGAAGACCTGTTCTATC
AAAGTTCACTTGCTTCTTGGAATTATAACACCAATATTACTGAAGAGATGTCC
AAAAGTGAATAATGGCTGGGGACAAATGGTCTGCCTTTTTANAGGAACAGTC
CANACTTGNC

B5 - 09/563,516

INC4179553

CTTGGCTCACAGGGGACGATGTCAAGCNCTTCCTGGCTCCTTCTCAGCCTTGT
TGCTGTAAGTCTGCTCAGTCCACCATTGAGGAACAGGCCAAGACATTTTGG
ACAAGTTTAACCACGAAGCCGAAGACCTGTTCTATCAAAGTTCACTTGCTTCT
TGGAATTATAACACCAATATTACTGAAGAGAATGTCCAAAACATGAATAATG
CTGGGGACA

INC4179240

GATCTTGGCTCACAGGGGACGATGTCAAGCTCCTTCCTGGCTCCTTCTCAGCC
TTGTTGCTGTAAGTGGCTGCTCAGTCCACCATTGAGGAACAGGCCAAGACATT
TTTGGACAAGTTTAACCACGAAGCCGAAGACCTGTTCTATCAAAGTTCACTTG
CTTCTTGGAATTATAACACCAATATTACTGAAGAGAATGTCCAAAACATGAN
TNATGCTGGGGGACAAATGGTCTGCCTTTTTAAAAGGAACAGTCCCAC

B7 - 09/563,516

INC3696047

TCAGTGGATGAAAAAGTGGTGGGAGATGAAGCGAGANATAGTTGGGGTGGT
GGGNCCTGTGCCCCATGATGAAACATACTGTGACCCCGCATCTCTGTTCCATG
TTTCTAATGATTACTCATTCAATTCGATATTACACAAGGACCCTTTACCAATTCC
AGTTTCAAGAAGCACTTTGTCAAGCAGCTAAACATGAAGGCCCTCTGCACAA
ATGTGACATCTCAAACCTCTACAGAAGCTGGACAGAAACTGTTCAATATGCTG
AGGCTTGGAAAATCAGAA
CCC

INC3699373

GCTCACAGGGGACGATGTCAAGCTCTTCCTGGCTCCTTCTCAGCCTTGTTGCT
GTAAGTGTGCTCAGTCCACCATTGAGGAACAGGCCAAGACATTTTTGGACA
AGTTTAACCACGAAGCCGAAGACCTGTTCTATCAAAGTTCACTTGCTTCTTGG
AATTATAACACCAATATTACTGAAGAGAATGTCCAAAACATGAATAATGCTG
GGGACAAATGGTCTGCCTTTTTAAAGGAACAGTCNACACTTGCCCAAATGTA
TCCATACAAGAAATTCAGAATCTCACAGTCCAGCTTC

INC4050931

ACTGGTGTAGCTGCAAGGATTGAGAATGGCATGCATTAGCTCACTTTCATTTA
ATCCATTGTCAAGGATGACATGCTTTCTTCACAGTAACTCAGTTCAAGTACTA
TGGTGATTTGCCTACAGTGATGTTTGGAATCGATCATGCTTTCTTCAAGGTGA
CAGGTCTAAAGAGAGAAGAATCCAGGGAACAGGTAGAGGACATTGCTTTTTTC
ACTTCCAAGGTGCTTGATCAACATCTCCCTGACAACACAAAACCTAGAGCCAG
GGGCCTCCGTGAACTCCCA

INC4874914

ATTCAGTGGATGTGATCTTGNGCTCACAGGGGACGATGTCAAGCTCTTCCTGG
CTCCTTCTGCAGCNTTGTGCTNTAACTGCTGCTCNGTCCACCATTGAGGAAC
AGGCCAAGACATTTTTGGACAAGTTTAACCACGAAGCCGAAGACCTGTTCTA
TCAAAGTTCACCTTGCTTCTTGGAATTATAACACCAATATTACTGAAGAGAATG
TCCAAAACATGAATAATGCTGGGGACAAATGGTCTGCCTTTTA

INC4539057

AAANATCTATGTTTTNCTCTTGAGGTGATTTTGTTGTATGTAAATGTTANTTT
CNTGGTATAGANAATNTAAGATGATAAAGATATCATTAAATGTCAAAACTAT
GACTCTGTTGAGAAAAAAATTGTCCAAAGACAACATGGCCAAGGAGAGAG
CATCTTCATTGACATTGCTTTCAGTATTTATTTCTGTCTCTGGATTTGACTTCT
GTTCTGTTTCTTAATAAGGNTTTTGTATTAGAGTATATTAGG

INC5106567

AGAGAGCATCTTCATTGACATTGCTTTCAAGTATTTATTTCTGNCTCTGGATTT
GACTTCTGTTCTGTTTCTTAATAAGGATTNTGTATTAGAGTATATTANGNAAA
GTGTGNATNTGGNCTCACAGGCTGTTCAAGGATAATCTAAATGTAAATGTCT
GTNGAATTTCTGAAGTTGANAACAAGGATATNTCATTGGAGCAAGNGTTGGA
TCTTGNATGGAATATGGATGGATCACTTGTAAG

INC4599088

GGATTTGACTTCTGTTCTGTTTCTTAATAAGGATTTTGTATTAGAGTATATTAG
GGAAAGTGTGTATNTGGTCTCACAGGCTGTTTCAGGGATAATCTAAATGTAAA
TGTCTGTTGAATTTCTGAAGTTGANAACAAGGATATATCATTGGAGCAAGTGT
TGGATCTCGTATGGAATATGGATGGATCACTTGTAAGGACAGTGCCTGGGAA
CTGGTGTAGCTGCAAGGATTGAGAATGGCATGCATTAGCTC

C7 - 09/563,516

INC4598818

CTAAAATCAGCTCTTGGAGATAAAGCATATGAATGGAACGACAATGAAATGT
ACCTGTTCCGATCATCTGTTGCATATGCTATGAGGCAGTACTTTTTAAAAGTA
AAANATCAGATGATTCTTTTTGGGGAGGAGGATGTGCGAGTGGCTAATTTGA
AACCAAGAATCTCCTTTAATTTCTTTGTCACTGCACCTAAAAATGTGTCTGAT
ATCATTCTAGAACTGAAGTTGAAAAGGCCATCAGGC

D1 - 09/563,516

INC5811393

GGAAGGGCGACTTCAGGATCCTTATGTGCACAAAGGTGACAATGGACGACTT
CCTGACAGCTCATCATGAGATGGGGCATATCCAGTATGATATGGCATATGCT
GCACAACCTTTTCTGCTAAGAAATGGAGCTAATGAAGGATTCCATGAAGCTG
TTGGGGAAATCATGTCACTTTCTGCAGCCACACCTAAGCATTATAAAATCCATT
GGTCTTCTGTCACCCGATTTTCAAGAAGACAATGAAACAGAAATAAACTTCC
TGCTCAAACAAGCACTCACGATTGTTGGGACTCTG

D2 - 09/563,516

EST801672

ATTTACAGTACTGGAAAAGTTTGGCAACCCAAAGAACCCACAAGAATGCTTAT
TACTTGAGCCAGGATTGGATGAAATAATGGCGACAACACAGACTACAACTCT
AGGCTCTGGGCATGGGAGGGCTGGAGGGCTGAGGTTGGCAAGCAGCTGAGG
CCGTTGTATGAAGAGTATGTGGTCCTGAAAAACGAGATGGCAAGAGCAAACA
ATTATAACGACTATGGGGATTATTGGAGAGGGGACTATGAAGCAGAGGGAG
CAGATGGCTACAACTATAACCGTAACCAGTTGATTGAAGATGTAGAACGTAC
CTTCGCAGAGATCAAGCCATTGTATGAGCATCTTCATGCCTATGTGAGGAGG
AAGTTGATGGATACCTACCCTTCCTACATCAGCCCCACTGGATGCCTCCCTGC
CCATTTGCTTGGTGATATGTGGGGTAGATTTTGGACAAATCTGTACCCTTTGA
CTGTTCCCTTTGCACAGAAACCAAACATAGATGTTACTGATGCAATGATGAAT
CAGGGCTGGGATGCAGAAAGGATATGTCAAGAGG

D3 - 09/563,516

EST1065165

TTTGTGTAATCTAGCATTTATTGAGTGTCTGCTTTGTGCTCAGCACTGCTCAA
ACACTGTGAGCAAATACAAAAGATTTAAGAGACTGGGTACCCACTTCAGAGG
GTGAACATACAGTTGGAATTTCCATTCACTCCACAGTTAGAGAACAGTAGAA
ATGAGTTTCTATCAGGCATGCTCTGGGAGTTCACGGAGGCCCTGCTCTAGTT
TTGTGTTGTCAGGGAGATGTTGATCAAGCACCTTGGAAGTGAAAAAGCAATG
TCCTCTACCTGTTCCCTGGATTCTTCTCTCTTTAGACCTGTCACCTTGAAGAAA
GCATGATCGATTCCAAACATCACTGTAGGCAAATCACCATAGTACTTGAAC
GAGTTACTGTGAAGAAAGCATGTCATCCTTGACAATGGATTAAATGAAAGTG
AGCTAATGCATGC

D4 - 09/563,516

EST1039409

GTGTATTTGGTCTCACAGGCTGTTTCAGGGATAATCTAAATGTAAATGTCTGTC
GAATTTCTGAAGTTGAAAACAAGGATATATCATTGGAGCAAGTGTTGGATCT
TGTATGGAATATGGATGGATCACTTGTAAGGACAGTGCCTGGGAACTGGTGT
AGCTGCAAGGATTGAGAATGGCATGCATTAGCTCACTTTCATTTAATCCATTG
TCAAGGATGACATGCTTTCTTCACAGTAACTCAGTTCAAGTACTATGGTGATT
TGCCTACAGTGATGTTTGGGAATCGATCATGCTTTCTTCAAGGTGACAGGTCTA
AAGAGAGAAGAATCCAGGGAACAGGTAGAGGACATTGCTTTTTCACTTCCAA
GGTGCTTGATCAACATCTCCCTGACAACACAAAACCTAGAGCCAGGGGCTCCG
TGAATCCCAGAGCATGCCTGAT

D5 - 09/563,516

EST1061120

TTTGTGTAAATCTAGCATTATTGAGTGTCTGCTTTGTGCTCAGCACTGCTCAA
ACACTGTGAGCAAATACAAAAGATTTAAGAGACTGGGTACCCACTTCAGAGG
GTGAACATACAGTTGGAATTTCCATTCACTCCACAGTTAGAGAACAGTAGAA
ATGAGTTTCTATCAGGCATGCTCTGGGAGTTCACGGAGGCCCTGGCTCTAGT
TTTGTGTTGTCAGGGAGATGTTGATCAAGCACCTTGGAAGTGAAAAAGCAAT
GTCCTCTACCTGTTCCCTGGATTCTTCTCTCTTTAGACCTGTCACCTTGAAGAA
AGCATGATCGATTCCAAACATCACTGTAGGCAAATCACCATAGTACTTGAAC
TGAGTTACTGTGAAGAA

D6 - 09/563,516

EST1065133

CTTCACAGTAACTCAGTTCAAGTACTATGGTGATTTGCCTACAGTGATGTTTG
GAATCGATCATGCTTTCTTCAAGGTGACAGGTCTAAAGAGAGAAGAATCCAG
GGAACAGGTAGAGGACATTGCTTTTTCACTTCCAAGGTGCTTGATCAACATCT
CCCTGACAACACAAAACCTAGAGCCAGGGGCCTCCGTGAACTCCCAGAGCATG
CCTGATAGAACTCATTCTACTGTTCTCTAACTGTGGAGTGAATGGAAATTC
CAACTGTATGTTACCCCTCTGAAGTGGGTACCCAGTCT

D7 - 09/563,516

EST1513548

AGAGAACCCTTATGACTCGATGGACATTGGAAAAGGAGAAAGCAATGCAGG
ATTCCAAAACAGTGATGATGCTCAGACTTCCTTTTAGCAAAGCACTTGTCATC
TTCCTGTATGTAAATGCTAACTTCATAGTACACAAAATATGAGAGTATACACA
TGTCATTAGCTATCAAACTATGATCTGTACAGTAAACGATGTCCAAAGAGC
ATCAGACTTGAGTGGACATCTTCACTGACATTGCTTTCAGTATTTATTTCTGCC
TAAGGATTTGACATCTCTTCTGTTTATTAATAGAGATGTTTATCTTAGCATAA
AAGAGGGAAATGTGCCTTTGGCTCACAGTCTATCCAGGGTGATATGGTTGGG
TAACTGGAGTTAGAAGATGAGATGATGTCTCTTGGGGGCAAGTGTTGGCTTC
GGTGTGGCATCTGGGCTGTGAACTGGTGGGACTGTTGAGGTTGAGA

EST1774662

TTGATGCTCTTTGGACAACCTTTTACTGAACAGATCATAGTTTTGATGGGTAAT
GACATGTGTATACTCTCATATTTTGTGTACTATGAAGTTAGCATTTACATACA
GGAAGATGACAAGTGCTTTGCTAAAAGGAAGTCTGAGCATCATCACTGTTTT
GGAATCCTGCATTGCTTTCTCCTTTTCCAATGTCCATCGAGTCATAAGGGCTC
TCTTCTCTTTTTGTTCATTTTTCTTCTTTCGACCTTTGATCCCAGTGACAATCA
GGATGATGATGCCAACCCTACCAGTGCCATCACAAACACCAAAAATAATCAG
CCATATGGTGACAGGAGGCTGGTAAGGTGGCTCAAGTGTTGGGTGAATCCCC
AGAAACTCCAGGCTGTTATCATTCAGGCCAAAGACATCATTGATGCGGCCCC
GAGA

E2 - 09/563,516

EST1776031

CTCTTCTGTTTCTTCTTCTGCTTTTTTTTTCTTCTCTTCTCAGTGCCCAACCCAA
GTTCAAAGGCTGATGAGAGAGAAAACTCATGAAGAGATTTTACTCTAGGGA
AAGTTGCTCAGTGGATGGGATCTTGGCGCACGGGGAAAGATGTCCAGCTCCT
CCTGGCTCCTTCTCAGCCTTGTTGCTGTTACTACTGCTCAGTCCCTCACCGAG
GAAAATGCCAAGACATTTTTAAACAACCTTTAATCAGGAAGCTGAAGACCTGT
CTTATCAAAGTTCACCTTGCTTCTTGGAATTATAATACTAACATTACTGAAGAA
AATGCCCAAAAGATGAGTGAGGCTGCAGCCAAATGGTCTGCCTTTTATGAAG
AACAGTCTAAGACTGCCCAAAGTTTCTCACTACAAGAAATCCAGACTCCGAT
CATCAAGCGTCAACTACAGGCCCTTCAGCAAAGTGGGTCTTCAGCACTCTCA
GCAGACAAGAACAAACAGTTGAACACAATTCTGAACACCATGAGCACCATT
ACAGTACTGGAAAAGTTTGCAACCCAA

E3 - 09/563,516

EST1874540

TTTTTTTTTTTTTGATGCTCTTTGGACAACGTTTACTGAACAGATCATAGTTTTG
ATAGCTAATGACATGTGTATACTCTCATATTTTGTGTACTATGAAGTTAGCAT
TTACATACAGGAAGATGACAAGTGCTTTGCTAAAAGGAAGTCTGAGCATCAT
CACTGTTTTGGAATCCTGCATTGCTTTCTCCTTTTCCAATGTCCATCGAGTCAT
AAGGGCTCTCTTCTCTTTTTGTTTCATTTTTCTTCTTTTCGACCTTTGATCCCAGT
GACAATCAGGATGATGATGCCAACCCTACCAGTGCCATCACAACACCAAAA
ATAATCAGCCATATGGTGACAGGAGGCTGGTAAGGTGGCTCAAGTGTTGGGT
GAATCCCCAGAACTCCAGGCTGTTATCATTGAGGCCAAAGACATCATTGAT
GCGGCCCCGAGACATCCTGATGGCATCTTCAACTTCACTTCTAGGAATGACAT
CAGACACATTTTGGGGTGAGGTGACAAAGAAGTAGAAGGAGACTCTTGGTTT
CAAATCACTCACTCGTACATCTTCTCTAGAAAAGGAAGTGTCTG

EST1874465

TTTAACTTCATATTGGTCCAGCAGCTTGTTTACTGTTCTCTTCTGTTTCTTCTTC
TGCTTTTTTTTTCTTCTCTTCTCAGTGCCCAACCCAAGTTCAAAGGCTGATGAG
AGAGAAAAACTCATGAAGAGATTTTACTCTAGGGAAAGTTGCTCAGTGGATG
GGATCTTGGCGCACGGGGAAAGATGTCCAGCTCCTCCTGGCTCCTTCTCAGCC
TTGTTGCTGTTACTACTGCTCAGTCCCTCACCGAGGAAAATGCCAAGACATTT
TTAAACAACCTTTAATCAGGAAGCTGAAGACCTGTCTTATCAAAGTTCACTTGC
TTCTTGGAATTATAATACTAACATTACTGAAGAAAATGCCCAAAGATGAGT
GAGGCTGCAGCCAAATGGTCTGCCTTTTATGAAGAACAGTCTAAGACTGCCC
AAAGTTTCTCACTACAAGAAATCCAGACTCCGATCATCAAGCGTCAACTACA
GGCCCTT

EST2447415

TTTTTTTTTTTTTTGGGTAAATCTAGCATTTATTGAGTGTCTGCTTTGTGCTCAG
CACTGCTCAAACACTGTGAGCAAATACAAAAGATTTAAGAGACTGGGTACCC
ACTTCAGAGGGTGAACATACAGTTGGAATTTCCATTCACTCCACAGTTAGAG
AACAGTAGAAATGAGTTTCTATCAGGCATGCTCTGGGAGTTCACGGAGGCCC
CTGGCTCTAGTTTTGTGTTGTCAGGGAGATGTTGATCAAGCACCTTGGAAGTG
AAAAAGCAATGTCCTCTACCTGTTCCCTGGATTCTTCTCTCTTTAGACCTGTCA
CCTTGAAGAAAGCATGATCGATTCCAAACATCACTGTAGGCAAATCACCATA
GTACTTGAAGTGAAGTACTGTGAAGAAAGCATGTCATCCTTGACAATGGATT
AAATGAAAGTGAGCTAATGCATGCCA

E6 - 09/563,516

EST2484649

TTTTTTGTGTAAATCTAGCATTTATTGAGTGTCTGCTTTGTGCTCAGCACTGCT
CAAACACTGTGAGCAAATACAAAAGATTTAAGAGACTGGGTACCCACTTCAG
AGGGTGAACATACAGTTGGAATTTCCATTCACTCCACAGTTAGAGAACAGTA
GAAATGAGTTTCTATCAGGCATGC

EST1623533

GGGGGGAGGTGGTGTGCGATCCCTCAGCCTGGGAGTTCTACGACGGCGAGGA
CTTCAGGGTCAAGCAGTGTACCTCTGTGAACATGGAGGACTTGGTGATAGCG
CACCACGGAATGGGCCACATCCACTATTTTCATGCAGTACAAAGACTTACCCC
TGACTTTCCGGGAGAGTGCCAACCCTGGTTTTTCATNAAGCTATTCGAGATATA
ATGGCTCTCTCAGTGTCTACCCCCAAGCATCTATACAGTCTCAACCTGCTTAG
CACTAAGGGCACTGGCTACGAGTATGACATCAACTTTCTAATGAAGATGGCC
CTCGACAAGATCGCCTTTATCCCCTTCAGCTACCTCATCGACCAGTCGCGCTC
GAGGGTCTTTGATGGAAGCATCACCAAGGAGAACTATAACCAGGAGTCCTGA
CCCTCAGCTGGAAGTATCAGGCTCTTTACCCCCAGTGCCAACACTCCAAGTG
ACTTTTACCCAGGATCCAAGTTCCACGTTCTGCCAGCTCCCATACGTCAGCT
ACTCTGTCANCTTCATCATCAAGTTGCAGTTACACGAGGCGCTGTATCCCCAA
CAGGACACACGNGGTCCCTACACAAATTCTAAATCTACAATTCCAAGTAGGC
GGGAATCTCCTCGCGTATACGGAATCTGGNCTACA

EST2887953

GGGAAAGCTGGAGATCTGAGGTCGGCAAGCAGCTGAGGCCATTATATGAAG
AGTATGTGGTCTTGAAAAATGAGATGGCAAGAGCAAATCATTATGAGGACTA
TGGGGATTATTGGAGAGGAGACTATGAAGTAAATGGGGTAGATGGCTATGAC
TACAGCCGCGGCCAGTTGATTGAAGATGTGGAACATACCTTTGANGAGATTA
AACCATTATATGAACATCTTCATGCCTATGTGAGGGCAAANTTGATGAATGC
CTATCCTTCCTATATCAGTCCAATTGGATGCCTCCCNGCTCATTTGCTTGGTGA
TATGTGGGGTAGATTTTGGACAAAACGTACTCTT

F2 - 09/563,516

EST3014598

TTACAGGACTGGAAAAGTATGCAACCCATAGAACCCACAGGAATGCTTATTA
CTTGAGCCAGGATTGGATGAAATAATGGCGACAAGCACAGACTACAACTCTA
GGCTATGGGCATGGGAGGGCTGGAGGGCTGAGGTTGGC